CLEARINGHOUSE
10·10

A Guide for Integrated Sample Selection for the Medicaid and Other Quality Control (QC) Systems

April 1978

U.S. DEPARTMENT OF HEALTH, EDUCATION, AND WELFARE
Health Care Financing Administration
Medicaid Bureau

REPORTS
QA
276
.5
G85
1978

QA
276.6
.G85
1978

CLEARINGHOUSE

PREFACE

This manual outlines a procedure for integrating the Quality Control (QC) samples for the Medicaid, AFDC and Food Stamp QC programs. Such integration has been tested out in several States and has proven to:

- Significantly reduce the numbers of staff and associated costs required for the QC effort.

- Greatly improve the utility of the data collected from the QC review process.

The Medicaid Bureau hopes States will find this manual useful. Any State wishing to implement an integrated sampling approach or having any questions on this approach should contact its Regional Medicaid Quality Control Office.

CONTENTS

I. INTRODUCTION

A. Rationale for Integrated Sample Selection

There are three federally mandated but state administered quality control systems for public assistance programs. These are the AFDC, Medicaid, and Food Stamp Quality Control programs. The purpose of these quality control systems is to measure and reduce the frequency of benefit errors, which are benefits disbursed for ineligible recipients or incorrect benefit amounts for eligible recipients.

The three quality control systems have many similarities. Each is based on a monthly random sample of cases. The quality control sample cases are investigated thoroughly to determine if errors have been made in eligibility determination or the benefit amount for the review month. The quality control review also collects information about the causes of any errors found. This information is used for corrective action planning to eliminate the causes of benefit errors.

In the past, there has been little interaction between the three quality control systems. Each system had its own sample (unrelated to the other system samples). When a case was reviewed, only the benefit for the program in whose sample the case fell was reviewed even though the same case might have also been receiving benefits from another program. In most states, individual quality control reviewers conducted only one type of review (AFDC, Food Stamp or Medicaid).

Until recently, this was a reasonable approach, since there was not much overlap in the universes sampled for the three quality control systems. Before July 1976, AFDC cases receiving food stamps were excluded from the

Food Stamp Quality Control system. Thus, there was no overlap in the universes for the AFDC and Food Stamp Quality Control systems. The Food Stamp Quality Control system has now been expanded to include AFDC cases receiving Food Stamps. The Medicaid Eligibility Quality Control system did not, but now does include AFDC cases (as well as some other types of cases which were previously excluded, e.g. SSI cases). The Medicaid Quality Control system has now been modified to use a sample of eligible cases rather than a sample of paid claims. Figure 1 shows the relationship which exists between the universes for the three quality control systems.

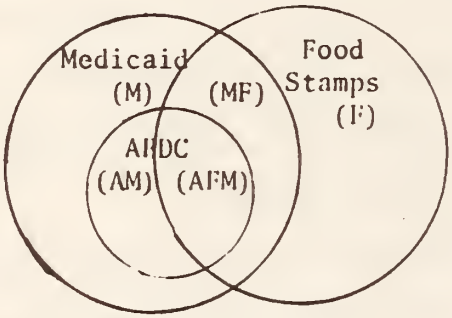


Figure 1. Overlapping universes for AFDC, Food Stamp, and Medicaid quality control.

Many families receive more than one of the benefits being considered. There would be a considerable degree of efficiency involved if a family could be included in more than one quality control sample. One home visit could replace two or three, thereby reducing travel, a significant cost in quality control. Since many of the items (e.g. assets and income) which have to be verified pertain to all three programs, and for other reasons discussed later, the cost of doing two reviews on one case is significantly less than the cost of doing two reviews on two different cases.

The term integrated sampling has been adopted to describe various methods whereby cases selected for one quality control sample can also be utilized as part of one (or two) of the other quality control samples. Several different approaches to integrated sampling are described in this guide. The difference among the various approaches are in:

1) The conditions required for implementation (e.g. some approaches require that the benefit combination in each case must be known while the sample is being selected, while other approaches do not).
2) The end result (some approaches produce self-weighting samples while the other approaches produce stratified samples which must be combined with weighting factors).
3) The relative efficiency (some approaches are less efficient because less multiple benefit cases are selected or not all of the benefits are reviewed in some of the multiple benefit cases).

It should be noted that there are economies involved in having generalized, or generic, quality control reviewers (reviewers who review cases in more than one program) even if integrated sample selection is not used. Since generic reviewers can perform more functions, each can be assigned to a smaller geographical area which reduces the total time and money spent travelling. Also, having cases to be reviewed located closer together reduces time spent waiting if time that would otherwise be spent waiting on a case can be used to do something on another case. Thus, a significant reduction in quality control costs may be achieved by using generic reviewers.

For the full potential efficiencies of integrated sampling to be realized, quality control reviewers must be able to complete any type of quality control review. The extra work that is involved in doing more than one review in some cases is more than offset by the economies involved in having generic workers. Thus, the percentage reduction in quality control review costs that can be achieved with generic reviewers and integrated sample selection is at least as large as the percentage reduction in the number of distinct cases to be reviewed.

With some of the integrated sampling procedures described later, the number of distinct cases to be reviewed can be reduced by about 50%, compared to the number of cases that would have to be reviewed if completely independent samples were used for AFDC, Medicaid, and Food Stamp Quality Control.

Since 50% of the overall quality control active case review costs is a significant amount of money in a state, integrated sample selection should be considered by states in spite of the additional complications involved.

B. History of Integrated Quality Control Sample Selection

The first opportunity and incentive for integrated quality control sample selection occurred in 1976 when the Food Stamp Quality Control system was expanded to include AFDC Food Stamp households, which had previously been excluded from Food Stamp Quality Control. The AFDC Food Stamp households were the first substantial group of cases to be sampled from for two quality control systems. Another effect of the 1976 changes was increased Food Stamp Quality Control sample sizes in many states. In some of the smaller states, the Food Stamp Quality Control sample size was increased by as much as 50%. At about the same time, quality control workloads were also increased by other changes, such as increased numbers of negative case action reviews.

Because of the increasing quality control sample sizes, there was considerable interest in methods to reduce quality control workloads. Many states submitted proposals to integrate sample selection for AFDC and Food Stamp Quality Control. Some of these were given tentative approval by the Department of Health, Education and Welfare and the Department of Agriculture. However, many questions about integrated sampling remained unanswered. To obtain answers to these questions, five States using integrated sample selection were visited, and their integrated sampling procedures studied. An Evaluation of Integrated Sample Selection for AFDC and Food Stamp Quality Control is a report on those visits. Some of the findings are summarized here.

Training specialized quality control reviewers to do other types of quality control reviews is not as difficult as had been anticipated. It can be accomplished with approximately one week of training and some additional supervision for the following several months. There was no indication of any reduction in the quality of either type of review when both types of reviews are done by all workers. Also, doing more than one type of review in the same case did not appear to reduce the quality of any of the reviews.

The only negative feedback on the integrated review process was concern with the duplication of effort involved in copying much of the same information onto the separate worksheets for each review in a multiple benefit case. An integrated worksheet, to collect all the information needed to fill out any of the quality control review schedues, has been developed and is being tested. If it proves effective it should greatly facilitate the integrated review process.

One of the concerns about integrated sampling is that the completion of reviews might be delayed. In some cases they are, but the delays are not long enough to cause problems in getting accurate information or meeting reporting deadlines.

The delays which do occur are caused by waiting for the Food Stamp participation lists so that when an AFDC or Medicaid case is assigned, it can be known whether it is a Food Stamp case also. The best way to avoid this is to draw the Food Stamp Quality Control sample from authorization lists, which are available during the review month. The reviews can be initiated promptly and cases can be dropped from Food Stamp Quality Control if they do not use their authorization to purchase food stamps.

When integrated sample selection is used, there is a strong argument for not initiating any of the reviews until the review month is over. The amount of Food Stamps and Medicaid benefits (if any) that a case uses during the review month could be influenced by either local agency or client knowledge of quality control sample selection during the review month. The only way to avoid this potential source of bias in the quality control findings would be to select the quality control samples in the month following the review month. (Because of this potential problem, this procedure may yet be mandated in Medicaid Quality Control for non-AFDC cases).

Integrated sample selection is more complicated than independent sample selection. However, if the complications are handled properly, the integrated samples are as statistically valid as independent samples. How much more complicated integrated sample selection is depends on the method used to obtain integrated samples. The first States to integrate quality control sample selection all used some variation of what is described later as the replacement method. The complications and frame requirements for this method were a barrier to the use of integrated sampling in some states. For that reason, other methods have been developed with less complications and frame requirements.

C. The Expanded Medicaid Quality Control System

The expanded Medicaid Quality Control System is designed to provide much more useful information than the previous MEQC system, which used relatively small samples compared to other quality control systems and covered only part of the total population of Medicaid recipients. The expanded Medicaid Quality Control sample sizes are several times as large as the previous MEQC sample sizes in most states. However, by integrating the Medicaid Quality Control sample selection with samples from other quality control systems, this expanded system can be implemented with only a small increase or even a decrease in the total quality control workload in a state.

The basic sampling design for the expanded Medicaid Quality Control system is a stratified sample. In States which have a 1634 contract with the Federal Government to determine eligibility for SSI cases, there are three strata as shown in figure 2.

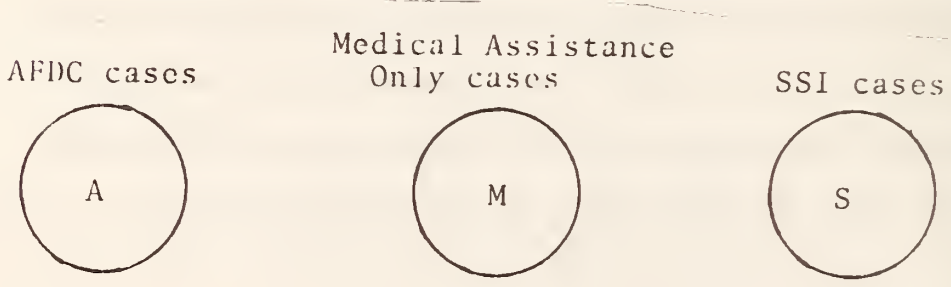


Figure 2. Strata for MQC in 1634 states.

The AFDC and SSI strata are already covered by other quality control systems (AFDC Quality Control and SSI Quality Assurance). The MQC sampling design involves using all of the AFDC Quality Control and SSI Quality Assurance sample cases as part of the Medicaid Quality Control sample. The rationale for this is that AFDC and SSI cases are automatically eligible for Medicaid. Therefore, a separate field investigation is not necessary to determine Medicaid eligibility if a definitive determination of AFDC or SSI eligibility has been made.

The number of Medicaid Quality Control sample cases to be selected from each stratum in each state is presented in the MQC Sampling Manual. In the AFDC and SSI strata, the number of MQC sample cases is determined by the AFDC Quality Control and SSI Quality Assurance sample sizes. The number of cases to be selected from the Medical Assistance Only strata is the same number of cases that were reviewed in the previous MEQC system.

In states which do not have a 1634 contract, there are only two separate strata in the MQC sampling design. The two strata are 1) AFDC cases, and 2) all other Medicaid cases. The SSI cases, which are included in the non-AFDC stratum, are not treated as a separate stratum because the state determines Medicaid eligibility for SSI cases in non-1634 states.

The number of MQC sample cases to be selected from the non-AFDC stratum in non-1634 states is the same as the number of cases reviewed in the previous MEQC system. Those SSI cases selected as part of the non-AFDC stratum may be replaced by cases reviewed by SSI-QA.

D. Additional Integration Possibilities

The basic Medicaid Quality Control sampling design uses integrated sample selection to some extent, but not to the full extent possible. Integrating Medicaid Quality Control sample selection with the sample selections for AFDC Quality Control and also with SSI Quality Assurance made it possible to enlarge the Medicaid Quality Control sample by about four times (from 17,575 to 77,410) without increasing the number of quality control cases with field investigations.

By taking advantage of additional integration possibilities among AFDC, Food Stamp, and Medicaid Quality Control, a state can implement the expanded Medicaid Quality Control system and end up reviewing less cases than are currently being reviewed.

Figure 3 is a schematic representation of the quality control sample selections assumed in the basic MQC sampling design.

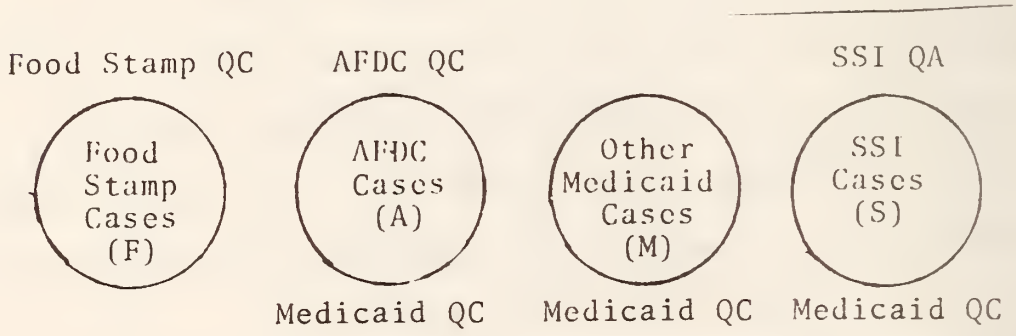


Figure 3. Cases included in various QC systems.

The additional integration that is possible is shown in figure 4.

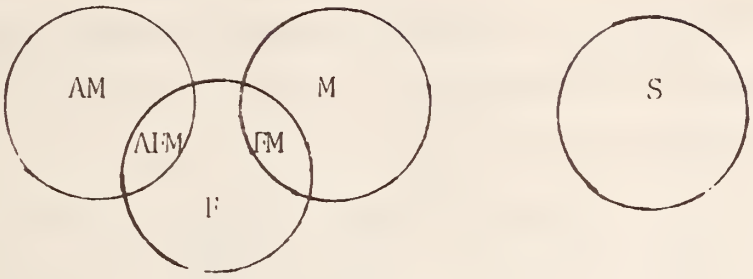


Figure 4. Additional integration possibilities.

Essentially, Food Stamp QC can now be integrated with both AFDC-QC and the Medical Assistance (non-AFDC) part of Medicaid QC.

In 1634 states, there would be no advantage to integrating the SSI Quality Assurance cases with any of the quality control systems except Medicaid, since states do not have to do field investigations in these cases.

There are many different methods which can be used to obtain integrated samples from overlapping universes. Each method has its advantages and disadvantages. The process of selecting the integrated sampling method to use would involve, first of all, eliminating those which could not be implemented because of some barrier such as the lack of required sample frames. Then, if more than one method could be implemented, priorities must be determined for the other features. For example, which is more important, efficiency or self-weighting samples.

The following sections describe several methods for selecting integrated samples, examples of each, and the sampling frame and weighting requirements. The information presented should enable a state to understand each method.

II. REPLACEMENT METHODS OF SELECTING INTEGRATED SAMPLES

A. Introduction

Replacement methods of selecting integrated samples use stratified samples from overlapping universes. Consider, for example, the two overlapping universes in figure 5.

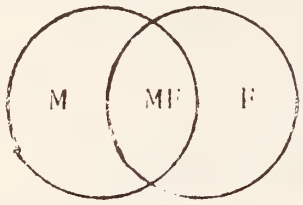


Figure 5. Two overlapping universes.

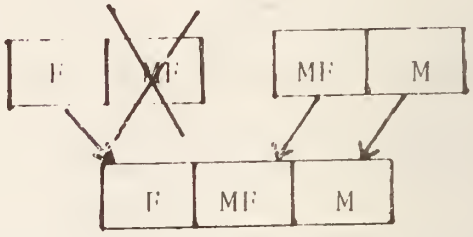


Figure 6. Stratified sample from each universe.

Each of the two universes consists of two strata, the overlapping part (MF) and the non-overlapping part (M or F). A stratified sample is obtained from each of the two universes, as shown in figure 6. One of the two samples from the overlapping stratum MF is deleted (or not even selected) and is replaced by some or all of the cases in the other sample from stratum MF. If the deleted sample is replaced by all of the other sample from the overlapping stratum, then one of the resulting stratified samples (F + MF in this example) is not a proportional stratified sample - which means that weighting must be used in the analysis of that sample.

Self-weighting samples for both universes can be obtained by replacing the smaller sample from MF with a random subsample of the same size from the other sample from MF.

The basic Medicaid Quality Control sampling design can be viewed as the non-proportional replacement method. In 1634 states, the Medicaid eligibility file is stratified as shown in figure 7. If a sample were taken from the entire Medicaid eligibility file, there would be some AFDC cases, some SSI cases, and some other Medicaid cases. Integrated sample selection is accomplished by replacing the AFDC cases with the cases from the AFDC Quality Control sample and replacing the SSI cases with cases from the SSI Quality Assurance sample.

Figure 7. Strata for MQC sample selection in 1634 states.

In non-1634 states, there are only two strata for MQC sample selection. These are 1) AFDC cases and 2) all other Medicaid cases. Sample selection is integrated with the AFDC Quality Control sample, by using the entire AFDC Quality Control sample in place of the AFDC cases that would otherwise have been obtained by sampling from the entire Medicaid eligibility file, and with SSI by substituting, for SSI cases falling into the sample, SSI-QA cases determined eligibility for Medicaid.

The replacement method can also be used to obtain integrated samples for all three state operated quality control systems (AFDC, Food Stamps, and Medicaid). The overlapping universes for these three quality control systems are shown in figure 8.

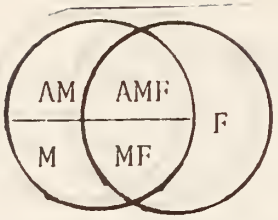


Figure 8. Overlapping universes for AFDC, Food Stamp, and Medicaid quality control.

The Medicaid universe to be sampled from consists of four strata, or types of cases. The two non-AFDC strata (M & MF) contain the types of non-AFDC cases which are to be selected for field investigations. In 1634 states, SSI cases are excluded. In non-1634 states, SSI cases are included in the two non-AFDC strata M & MF).

The replacement method sampling procedure for integrating the Medicaid, AFDC, and Food Stamp Quality Control is presented in the next section.

B. Sampling Procedure

1. Select the AFDC Quality Control sample in the normal manner using the interval

$$I_a = \frac{\text{estimated average caseload size}}{\text{required AFDC sample size (allowing for drops)}} \times 6$$

All of these cases will be included in the MQC sample. All (or part) of the cases which receive Food Stamps will be included in the Food Stamp Quality Control sample.

2. Select the sample of non-AFDC Medicaid cases* to be reviewed using the interval

$$I_m = \frac{\text{estimated average caseload size*}}{\text{required non-AFDC MQC sample size (allowing for drops)}} \times 6$$

*This is the number of Medical Assistance Only cases in 1634 states and the number of non-AFDC Medicaid cases (including SSI cases) in non-1634 states. Some or all of these will be included in the Food Stamp Quality Control sample.

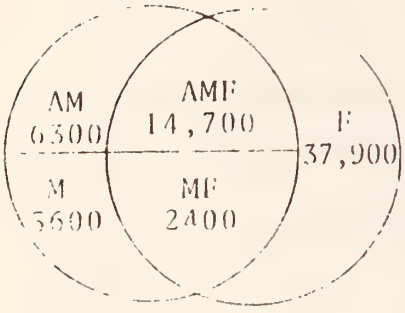
3. Select a sample of Food Stamp only cases (excluding any Food Stamp cases which are also receiving AFDC or are Medicaid beneficiaries) using the interval

$$I_f = \frac{\text{estimated average Food Stamp only caseload size}}{\text{required Food Stamp only sample size (allowing for drops)}} \times 6$$

The size of the Food Stamp only sample is the **required Food Stamp** sample size minus the number of Food Stamp cases obtained in steps 1 & 2.

C. Example

This example illustrates the use of the replacement method in a 1634 state with caseload configuration shown in figure 9. SSI cases have been excluded from the Medicaid universe.



	Caseload	Required Sample
Food Stamp Caseload (total	55,000	1,200
Non-cash Medicaid Caseload (total)	8,000	175
AFDC Caseload	21,000	550

Figure 9. Overlapping universes with SSI cases not included in the Medicaid universe.

The Food Stamp sampling interval will be

$$I_f = \frac{55,000}{(1.1)(1200)} \times 6 = 41.67 \times 6 = 250$$

[allowing for a 10% drop rate]

The number of Food Stamp cases selected for a six month period will be about 1320. Of these about

$$1320 \times \frac{2400}{55,000} = 58$$

will be from strata MF. These will be replaced by that number (58) of MF cases obtained while selecting the Medicaid Quality Control sample.

There will be about $1320 \times \frac{14,700}{55,000} = 353$ AMF cases in the Food Stamp sample. These will be replaced by that number (353) of AMF cases obtained while selecting the AFDC sample.

The AFDC interval will be

$$I_a = \frac{21,000}{(1.1)(550)} \times 6 = \frac{21,000}{605} \times 6 = 208$$

[allowing for a 10% drop rate]

The 605 AFDC sample cases will provide about

$$605 \times \frac{14,700}{21,000} = 423 \text{ AMF cases.}$$

To obtain a self-weighting Food Stamp sample, only a random sample of 353 of these should be selected for Food Stamp reviews.

The Medicaid interval will be

$$I_m = \frac{8000}{(1.1)(175)} \times 6 = 249 \qquad \left[\text{allowing for a 10\% drop rate}\right]$$

Of the 192 cases that will be selected, about $192 \times \frac{2400}{8000} = 58$ will be MF cases. This is exactly the number needed to replace the 58 MF cases in the Food Stamp sample.

The table below shows the number of cases and reviews for each benefit combination. The total number of quality control reviews is 2922. There are only 1706 cases selected for field investigation which is a 42% reduction due to integrated sample selection.

Benefit Combination	Cases Selected	AFDC Reviews	Medicaid Reviews	Food Stamp Reviews
F	909			909
MF	58		58	58
AMF	423	423	423	353
AM	182	182	182	
M	134		134	
	1706	605	797	1320
			+200 SSI	
			997	

In this example, disproportionate replacement would have resulted in 70 less Food Stamp only cases to be reviewed. This additional 2% reduction in cases to be reviewed would probably not be worth the loss of a self-weighting Food Stamp Quality Control sample. In other caseload configurations disproportionate replacement might be considerably more efficient.

D. Frame Requirements

For this method to be implemented, the following conditions should be met.

1) It must be possible to identify (at the time of sample selection) those AFDC Quality Control sample cases which also receive Food Stamps.

2) The fraome used to select the Food Stamp Quality Control sample cases must specify which are AFDC cases and which are non-AFDC cases and which are non-AFDC Medicaid cases.

3) The frame used to select Medicaid cases must specify which cases are AFDC, and in 1634 states, which cases are SSI and which cases are neither of these. It must also specify which cases receive Food Stamps.

E. Weighting Required

This method results in the following weighting requirements.

1) No weighting is required in the compilation of AFDC Quality Control findings.

2) Weighting is always required in the compilation of Medicaid Quality Control findings and is required for Food Stamp Quality Control if disproportionate replacement is used. Weighting methods are explained in Part E of Section IV.

III. SOCIAL SECURITY NUMBER METHOD

A. Introduction

This method was first used by the South Dakota Department of Social Services to integrate sample selection for AFDC and Food Stamp Quality Control. This method is of special interest because it produces self-weighting samples without requiring that the benefit combination of cases be indicated in the sampling frames. This method assumes that the last four digits of the Social Security numbers in a given state caseload are random.

B. Sampling Procedure

Each month some method is used to generate a pre-specified number of four digit random numbers. The number of random numbers needed to select the AFDC Quality Control sample may be different from the number of random numbers needed to select the Food Stamp sample. Let R_a, R_f, and R_m be the respective numbers of random numbers used for the three samples.

The AFDC Quality Control sample is obtained by comparing the last four digits of the Social Security number of the payee in each AFDC case to all R_a random numbers being used to select the AFDC Quality Control sample. If there is a match, the case is selected to be included in the AFDC Quality Control sample. The probability of selection for any AFDC case is $R_a/10,000$. R_a is determined as follows.

$$R_a = \frac{\text{Average Monthly Sample Size}}{\text{Estimated Average Caseload Size}} \times 10,000$$

If the estimated average AFDC caseload size is 21,000 and the average monthly sample size is 550/6 = 92, then

$$R_a = \frac{92 \times 10,000}{21,000} = 44$$

The Food Stamp Quality Control sample is selected in the same way, using the Social Security number of the Food Stamp household head. R_f, the number of random numbers needed to select the Food Stamp sample, is

$$R_f = \frac{\text{Average Monthly Sample Size}}{\text{Estimated Average Caseload Size}} \times 10,000$$

If the estimated average Food Stamp caseload is 55,000 and the average monthly sample size needed is 200, then

$$R_f = \frac{200 \times 10,000}{55,000} = 37$$

To obtain an integrated sample, 37 of the same random numbers used to select the AFDC sample must be used.

The Medicaid Quality Control sample is selected in the same way, using the Social Security number of the primary person in the Medicaid case. R_m, the number of random numbers needed to select the non-AFDC Medicaid sample is

$$R_m = \frac{\text{Average Monthly Sample Size}}{\text{Estimated Average Caseload Size}} \times 10{,}000$$

If the estimated average caseload size is 8,000 and the average monthly non-AFDC Medicaid sample size is 175/6, then

$$R_m = \frac{(175/6) \times 10{,}000}{8{,}000} = 36$$

To obtain an integrated sample, 36 of the random numbers used to select AFDC and Food Stamp cases are used to select the Medicaid sample.

C. Example

The table below shows the number of cases of each type in the quality control samples in the example state.

The total number of cases with field investigations is 1571, which is 42% less than the total number of reviews (2704).

Benefit Combination	Caseload Size	AFDC Sample Cases	Medicaid Sample Cases	Food Stamp Sample Cases
AMF	14,700	388	388	326
AM	6,300	166	166	
MF	2,400		53	53
M	5,600		123	
F	37,900			840
		554	730	1220
SSI			200	
			930	

D. Frame Requirements

The only frame requirement for this method is that the frames for selecting the three quality control samples all contain a Social Security number for each case. The statistical validity of this method is not affected by an occasional case with no Social Security number if a provision is made to generate random numbers for the missing Social Security numbers. This will, however, reduce the amount of overlap in the samples and thus reduce the efficiency of the method.

E. Weighting Required

The Social Security number method produces self-weighting samples for the AFDC and Food Stamp Quality Control systems and the normal stratified sample for the Medicaid Quality Control system.

IV. ADDITION METHOD OF SELECTING INTEGRATED SAMPLES

A. Introduction

The addition method of selecting integrated samples from two overlapping universes is shown in figure 10.

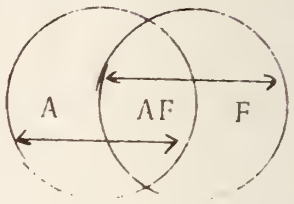


Figure 10.

Addition method of obtaining integrated samples.

The sampling procedure is to obtain a simple (or systematic) random sample from each of the two universes and to review all of the benefits in all of the cases selected. This is called the addition method because the cases in the second sample from the overlap domain are added to (rather than being replaced by) the cases in the first sample from the overlap domain.

The primary advantage of this method is that each universe can be sampled from independently. That is, it is not necessary to know the benefit combination at the time of sample selection. Another advantage of this method is that a greater proportion of the sample cases will be multiple benefit cases and the total number of distinct cases will be less than would be obtained with the replacement method. The primary disadvantage of this method is that data analysis and reporting for all of the resulting samples will require weighting.

B. Sampling Intervals

If the total number of cases to be reviewed for each quality control system is a specified number, the sampling intervals to be used for the additional method of obtaining integrated samples are determined as follows. The overlapping universes for AFDC, Food Stamp, and Medicaid Quality Control are shown in Figure 11. As before, SSI cases are excluded from the non-AFDC Medicaid universe in 1634 states and included in other states.

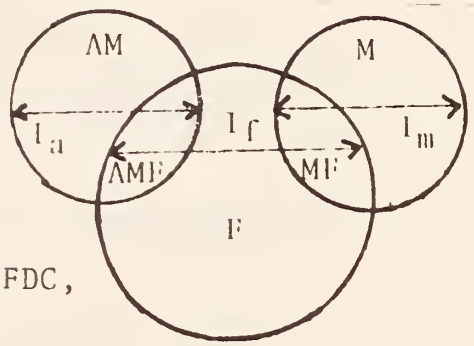


Figure 11.

Overlapping universes for Medicaid, AFDC, and Food Stamp Quality Control sample selection.

Applying intervals I_m, I_f, and I_a to their respective universes will produce samples of sizes Q_m, Q_r, and Q_a. Let S_m be the required number of Medicaid cases to be selected (excluding AFDC cases in all states, and also excluding SSI cases in 1634 states). Let S_a and S_f be the required numbers of cases to be selected for AFDC and Food Stamp quality control and N_j be the number of cases in stratum j. Then

$$S_m = Q_m + Q_f \frac{N_{mf}}{N_{mf} + N_m}$$

$$= Q_m + Q_f P_{fm} \qquad \text{where } P_{fm} = \frac{N_{mf}}{N_{mf} + N_m}$$

$$S_a = Q_a + Q_f \frac{N_{amf}}{N_{amf} + N_f + N_{mf}}$$

$$= Q_a + Q_f P_{af} \qquad \text{where } P_{af} = \frac{N_{amf}}{N_{amf} + N_f + N_{mf}}$$

$$S_f = Q_f + Q_a \frac{N_{amf}}{N_{amf} + N_{am}} + Q_m \frac{N_{mf}}{N_{mf} + N_m}$$

$$= Q_f + Q_a P_{fa} + Q_m P_{fm} \qquad \text{where } P_{fa} = \frac{N_{amf}}{N_{amf} + N_{am}}$$

Solving these three equations for the three unknowns (Q_m, Q_a, Q_f) gives

$$Q_m = S_m - Q_f P_{mf}$$

$$Q_a = S_a - Q_f P_{af}$$

$$Q_f = \frac{S_f - S_a P_{fa} - S_m P_{fm}}{1 - P_{af} P_{fa} - P_{mf} P_{fm}}$$

The sampling intervals are then computed by dividing each universe size by the number of cases to be selected.

$$I_m = \frac{N_m + N_{mf}}{Q_m} \times 6$$

$$I_a = \frac{N_{am} + N_{amf}}{Q_a} \times 6$$

$$I_f = \frac{N_{amf} + N_{mf} + N_f}{Q_f} \times 6$$

States should oversample to some degree (i.e. reduce the sampling interval) to insure that minimum sample size requirements are met in all programs.

C. Example

The example in the previous section assumed the following caseload configuration and required sample sizes.

$N_m = 2,400$ $S_m = 192$

$N_{mf} = 5,600$

$N_{amf} = 14,700$ $S_a = 605$

$N_{am} = 6,300$

$N_f = 37,900$ $S_f = 1,320$

If the addition method were used in this same situation the following results would be obtained. The number of cases selected from each of the frames would be

$$Q_f = \frac{S_f - S_a P_{fa} - S_m P_{fm}}{1 - P_{af} P_{fa} - P_{mf} P_{fm}}$$

$$= \frac{1320 - 605\ (.70) - 192(.30)}{1 - (.27)(.70) - (.044)(.30)} = 1052$$

$$Q_m = S_m - Q_f P_{mf} = 192 - 1052(.044) = 146$$

$$Q_a = S_a - Q_f P_{af} = 605 - 1052(.27) = 321$$

The number of sample cases with each benefit combination would be

$M_f = 722$ $M_{am} = 96$ $M_{amf} = 508$

$M_{fm} = 90$ $M_m = 102$

There would thus be 1320 cases selected for Food Stamp reviews, 605 cases selected for AFDC reviews, and 192 cases (in addition to the 605 AFDC cases) selected for Medicaid reviews. The total number of reviews is 1320 + 605 + 605 + 192 + 200 = 2922. The number of distinct cases with field investigations is 722 + 90 + 508 + 96 + 102 = 1518. The percentage reduction in cases with field investigations is

$$\frac{2922 - 1518}{2922} = 48\%$$

D. Frame Requirements

This is the only method which does not have any sample frame requirements in addition to the frames that would be required to select independent samples for each of the quality control systems. However, estimates of the percentage of multiple benefit cases are necessary.

M. Weighting Required

Weighting is required for all three quality control samples.

For all of the methods of selecting integrated samples described here, the following procedure can be used whenever weighting is required. The formula for computing any error rate is

$$\hat{P} = \sum_h W_h \hat{P}_h$$

Where $\sum_h$ is the sum over all strata covered by the QC system that the error rate is for.

W_h = Size of caseload in stratum h / QC system universe size

$\hat{P}_h$ is the error rate for all QC sample cases from stratum h

Consider, for example, the following caseload configuration and error rates.

Benefit Combination	Cases in Universe	Cases in Sample	Error Rate
F	37,900	909	20%
MF	2,400	58	22%
AMF	14,700	423	12%
AM	6,300	182	10%
M	5,600	134	14%
SSI	42,000	200	2%

The error rates for AFDC, Food Stamp, and Medicaid quality control would be

$$P_a = \frac{14,700}{21,000} \times 12 + \frac{6,300}{21,000} \times 10 = 11.4\%$$

$$P_f = \frac{37,900}{55,000} \times 20 + \frac{2,400}{55,000} \times 22 + \frac{14,700}{55,000} \times 12 = 17.95\%$$

$$P_m = \frac{2,400}{71,000} \times 22 + \frac{14,000}{71,000} \times 12 + \frac{6,300}{71,000} \times 10 + \frac{5,600}{71,000} \times 14 + \frac{42,000}{71,000} \times 2 = 6.39\%$$

V. OPTIMUM ALLOCATION METHOD

A. Introduction

The previous methods all result in the selection of different numbers of cases with each benefit combination. None of the allocations are optimal in the sense of minimizing the variance of error rate estimates. The optimal number of cases with a particular benefit combination to be selected can be computed once it is decided what to minimize and what the constraints are. An example of an optimum allocation of cases based on one set of assumptions is described in Appendix A. Note that constraint equations for minimum sample sizes would have to be included if this method was used.

B. Sampling Procedure

For an optimal allocation design, each case would ideally be exposed to sampling only once and would be selected with a probability which would be the same for all cases with the same benefit combination and different for cases with different benefit combinations.

C. Frame Requirements

To select cases in the most optimal manner, it would be necessary to have a single frame (computer file or listing) of the combined AFDC, Food Stamp and Medicaid caseloads. Each case should be listed only once and the benefit combination should be indicated. Unless a computer file of this type already exists, it would probably not be worth the effort to create such a frame just for the purpose of integrated sample selection.

D. Weighting Requirements

Any optimized sample will require weighting for reporting and data analysis in all three quality control samples.

E. Relative Efficiency

The measure of efficiency that has been used here to compare the various methods of obtaining integrated samples is the number of cases reviewed divided by the number of reviews obtained. Optimum allocation may or may not be more efficient in this respect than the addition method. The efficiency of the optimum allocation method will depend on the constraints that are set and the variable chosen (variance, cost, sample size) for optimization. Allocation methods can be deveoped to minimize total reviews while meeting minimum requirements.

The table below shows the number of cases with each benefit combination that would be selected in the example state according to the optimal allocation in Appendix A, an allocation which minimizes costs subject to keeping variances at or below current program requirements without regard to minimum sample sizes. The total number of reviews is 2960. There is a 41% reduction in the number of cases to be reviewed with this method of integrated sample selection.

Benefit Combination	Cases Selected	AFDC Reviews	Medicaid Reviews	Food Stamp Reviews
F	574			574
MF	262		262	262
AMF	401	401	401	401
AM	155	155	155	
M	348		348	
TOTAL	1740	556	1166	1238

APPENDIX A

An Example of Optimal Integrated Samples
Rudolph E. Schwartz, Office of Research and Statistics
Social Security Administration, Woodlawn, MD. 21235

An optimal sample takes into consideration three elements:

(1) Reliability or allowable sampling error.

(2) Variability of the universe or sub-universes from which samples are drawn.

(3) Cost of sampling which considers setup, travel, interview and verification costs.

The difficulty with designing an optimal sample is that each of the various programs has its own values for the three elements mentioned above. Thus, the sample that is optimal for one program may not be optimal for another.

In order to solve such a problem for stratified samples and several programs one must solve a nonlinear programming problem.

Mathematical Formulation of an Optimal Stratified Integrated Sampling Problem Samples

$$\text{MIN } C = \sum_h a_h n_h \qquad h = 1 \ldots L,\ j = 1 \ldots K$$

Subject to

$$V_j \leq K_j = \sum_{h=1}^{L} \frac{W_h^2 S^2_j}{n_h} \text{ (neglecting finite correction factor)}$$

$$j = 1 \ldots K,\ 0 \leq n_h \leq N_h$$

where N = Total units in population
N_h = Total units in h stratum
n_h = Sample units in hth stratum
W_h = N_h/N = Stratum weight
S_h^2 = Estimate of population variance in h stratum

The next page gives an example of a typical optimal solution which has been solved by a nonlinear program package. Since many of the States do not have access to such a package, we will be happy to solve a States design problem if they will furnish us with the required data.

Example - Design of an optimal integrated sample. Assume the present three programs have the following sample size requirements.

Program	Sample	SRS Variance* (10^{-4})
Aid for Dependent Children (AFDC)	800	1.10
Food Stamps (FS)	1,200	0.61
Medicaid (MED)	275	2.76

*SRS Variance = Simple random sample variance equivalent for sample sizes if P = .095 (QC error rate)

Construct mutually exclusive strata by merging the files of the three programs into seven mutually exclusive strata.

Stratum No.		Population Size (000)	Factor	Cost Sampling	Unit Cost
1	AFDC	9	1	9.50	9.50
2	AFDC, FS	18	1/2	11.50	5.75
3	AFDC, FS, MED	2	1/3	13.50	4.50
4	AFDC, MED	1	1/2	11.50	5.75
5	FS	23	1	9.50	9.50
6	FS, MED	7	1/2	11.50	5.75
7	MED	45	1	9.50	9.50

MINIMIZE:

$$C = 9.50n_1 + 11.50n_2 + 13.50n_3 + 11.50n_4 + 9.50n_5 + 11.50n_6 + 9.50n_7$$

SUBJECT TO:

$$0 \leq n_1 \leq N_1 \cdots ; \; 0 \leq n_7 \leq N_7$$

$$\left(\frac{9}{30}\right)^2 \frac{(P_{11})(1-P_{11})}{n_1} + \left(\frac{18}{30}\right)^2 \frac{(P_{21})(1-P_{21})}{n_2} + \left(\frac{2}{30}\right)^2 \frac{(P_{31})(1-P_{31})}{n_3} + \left(\frac{1}{30}\right)^2 \frac{(P_{41})(1-P_{41})}{n_4} \leq 1.10 \times 10^{-4}$$

$$\left(\frac{18}{50}\right)^2 \frac{(P_{22})(1-P_{22})}{n_2} + \left(\frac{2}{50}\right)^2 \frac{(P_{32})(1-P_{32})}{n_3} + \left(\frac{23}{50}\right)^2 \frac{(P_{52})(1-P_{52})}{n_5} + \left(\frac{7}{50}\right)^2 \frac{(P_{62})(1-P_{62})}{n_6} \leq 0.61 \times 10^{-4}$$

$$\left(\frac{2}{55}\right)^2 \frac{(P_{33})(1-P_{33})}{n_3} + \left(\frac{1}{55}\right)^2 \frac{(P_{43})(1-P_{43})}{n_4} + \left(\frac{7}{55}\right)^2 \frac{(P_{63})(1-P_{63})}{n_6} + \left(\frac{45}{55}\right)^2 \frac{(P_{73})(1-P_{73})}{n_7} \leq 2.76 \times 10^{-4}$$

Where $P_{11} = 09$, $P_{21} = .10$, $P_{31} = 08$, $P_{41} = 07$

$P_{22} = 06$, $P_{32} = 09$, $P_{52} = .10$, $P_{62} = 08$

$P_{33} = 08$, $P_{43} = .10$, $P_{63} = 07$, $P_{73} = .10$

$n_1 = 205$, $n_2 = 504$, $n_3 = 54$, $n_4 = 19$, $n_5 = 564$, $n_6 = 144$, $n_7 = 226$

Cost = 17852 HRS

CMS LIBRARY
3 8095 00014181 8